324-77 CROSS REFERENCE SEARCH ROOM

May 17, 1955 M. V. KALFAIAN 2,708,688

PHONETIC PRINTER OF SPOKEN WORDS

Filed Jan. 25, 1952 3 Sheets-Sheet 1

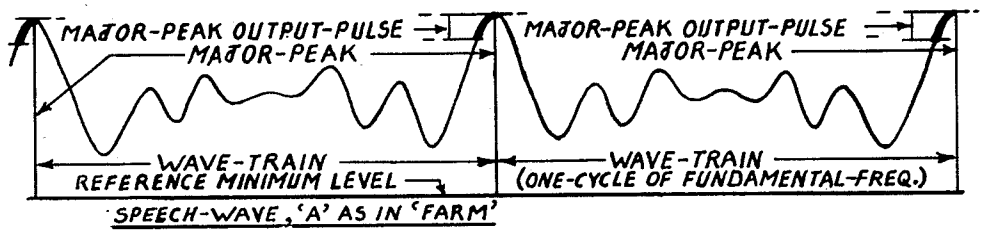


Fig. 1

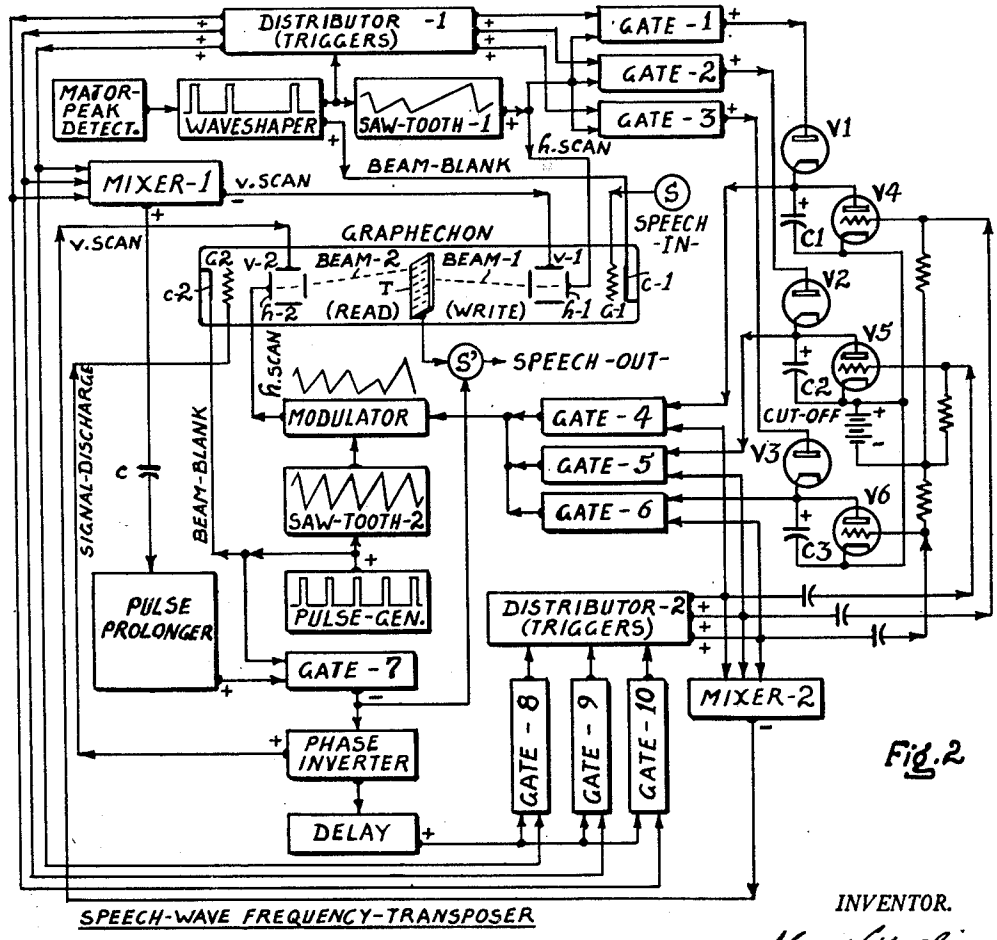


Fig. 2

INVENTOR.

Myron V. Kalfaian

May 17, 1955 M. V. KALFAIAN 2,708,688

PHONETIC PRINTER OF SPOKEN WORDS

Filed Jan. 25, 1952 3 Sheets-Sheet 2

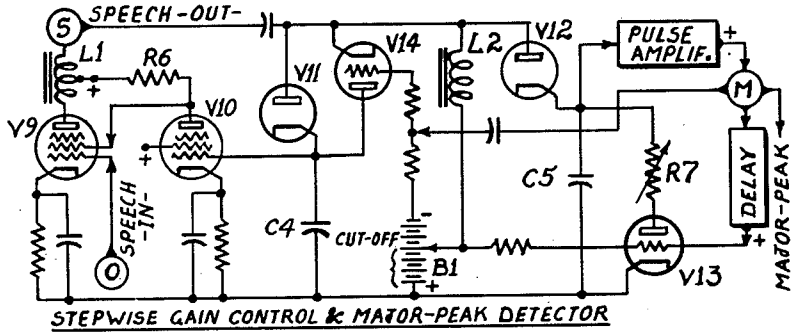


Fig. 3

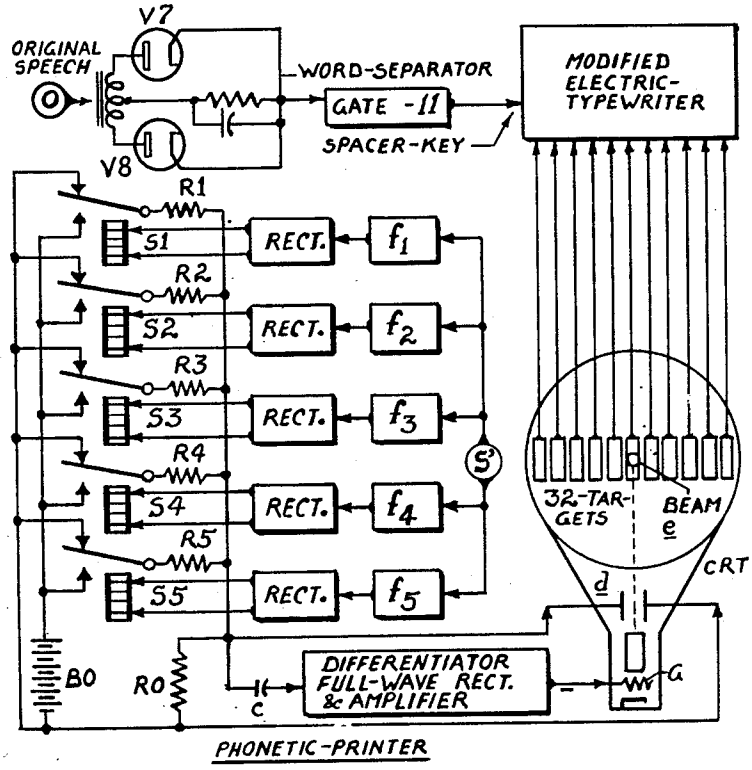


Fig. 4

INVENTOR.

Myron V. Kalfaian

May 17, 1955 M. V. KALFAIAN 2,708,688

PHONETIC PRINTER OF SPOKEN WORDS

Filed Jan. 25, 1952 3 Sheets-Sheet 3

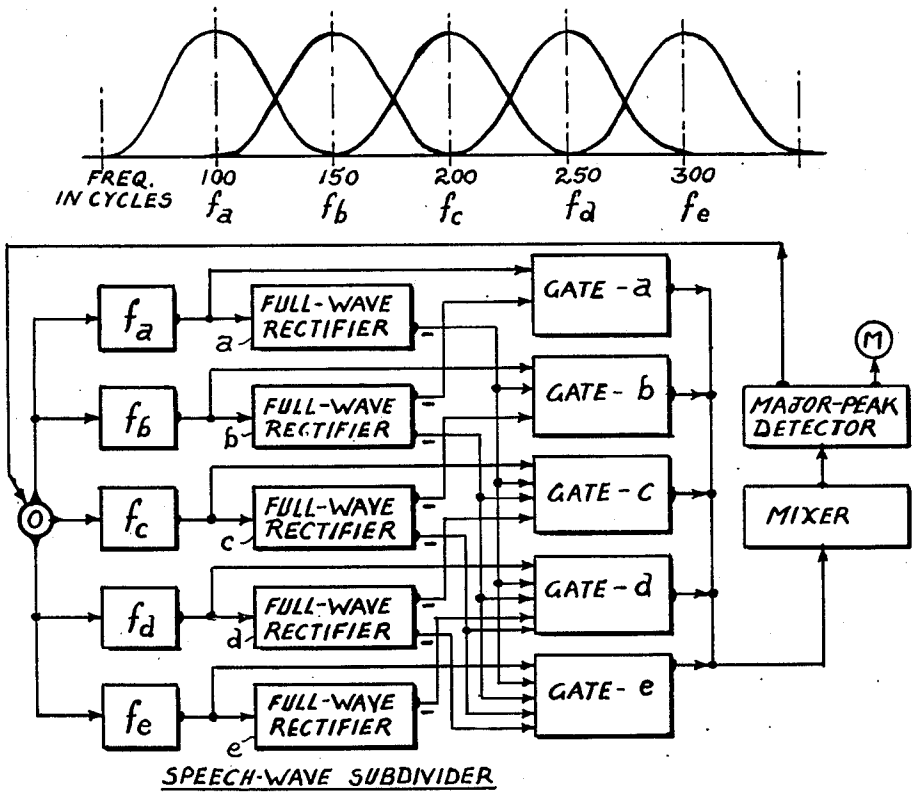


Fig. 5

INVENTOR.

Melgun V. Kalfaian

United States Patent Office 2,708,688 Patented May 17, 1955

2,708,688

PHONETIC PRINTER OF SPOKEN WORDS

Meguer V. Kalfaian, Los Angeles, Calif.

Application January 25, 1952, Serial No. 268,243

15 Claims. (Cl. 178—31)

This invention relates to the analysis of speech waves, and more particularly to those waves that are responsible for the intelligibility of phonetic characters in spoken sounds. Its main object is to provide methods and means for the analysis of various wave-patterns during propagation of articulate sounds for the selection and control of phonetic characters contained therein. A corollary object is to provide methods and means to translate the selected phonated characters into discrete signals for the actuation of character-printing keys, for example, the keys of a modified electric typewriter. A further object is to provide methods and means of transposing all frequency components of importance contained in speech waves based on unknown fundamental (pitch) frequencies, to frequency regions based on a single known fundamental frequency, whereby eliminating most all unknown variables while selecting phonetic characters therefrom.

Basic theory of phonetic sounds

In ordinary speech, each phonetic sound is produced in substantially replica wave trains that repeat successively at a fundamental (pitch) frequency, during propagation of the articulated sound. The succession of these wave trains is effected by fairly regular puffs of air from the glottis, which are set into vibration in the momentarily formed resonant cavities of the vocal system. Each wave train contains all the phonetic information necessary, and its specific wave shape is formed by the number of frequency components that are produced in these cavities, and their relations with regard to frequency positions and relative amplitude levels one with another. The basic frequency components composing a pure phonetic sound are independent of characterization components, the latter of which are mainly produced by the larynx. The composite structure of these latter components is inconsistent in form, and it varies in a complex manner with the varying pitch of the speaker's voice. However, the frequency ratios of the basic components, and their relative amplitudes, remain substantially constant with respect to the fundamental frequency; even though the frequency locations of all these components may change in the entire spectrum band of the voice. Thus, the human intelligence interprets phonetic sounds by measuring the ratios of basic frequency components, and their relative amplitudes, with respect to the fundamental frequency, without regard to the characterization components; the latter of which is interpreted as a form of voice quality.

To define characterization complexities, as an example, let one speaker pronounce certain phonetic sound (in natural voice) at first and second fundamental frequencies. The listener can easily recognize the characteristic quality of the sound to be of the same speaker. But when the sound at the first fundamental is recorded and reproduced at the second fundamental (by speeding or retarding the movement of reproduction), the listener can easily detect the phonetic sound but cannot recognize the characteristic quality of the voice. The concept of this peculiar condition has been advanced by actual recordings of male and female voices on magnetic tape. The consonant sounds had been both preceded and succeeded by vowel sounds, and the reproduction speed had been varied randomly. In all cases, the phonetic sound had been recognizable by a group of listeners. Actual tests have also shown that the characterization frequency components are mostly outside the regions of the band embracing the basic components of each phonetic sound, the condition of which indicates one reason why the human intelligence can easily distinguish between the characterization and pure phonetic sounds. For practical speech wave analysis however, frequency standardization, during propagation of the speech wave, is essential, since the speaker's pitch varies from one instant to another, and thus by changing the time period of any wave pattern, the frequency positions of the information bearing peaks are shifted widely in the voice spectrum.

Mode of standardizing the variables

In order to eliminate the variables of pitch, for any voice, it is possible to measure the differences of the fundamentals with respect to a reference frequency, and shift all frequency components of the speech to frequency regions where they will all be based on a single reference fundamental frequency. The characterization frequency components extending beyond the bandwidth that embraces the total number of basic components composing all pure phonetic sounds may then be filtered out therefrom. Thus by rearranging the frequency positions of all the components during propagation of the speech waves, so that they will all be based on a single reference fundamental frequency, the frequency positions of the basic components can be standardized.

Extraction of fundamentals

In reference to the foregoing, as each puff of air enters the resonant cavities of the vocal tract, there forms a high-peaked surge, which produces the first constituent wave of a wave-pattern higher in amplitude than any of the other waves contained therein. Similarly, the puffs of air are in forward (pushing) direction, which indicates that these major peaks are always in the same direction. Thus by pre-polarizing the incoming waves, from microphone to the point of analysis, the major peaks can be detected to mark the arrival and ending of any wave pattern for subdividing the speech waves into one cycle portions of the fundamental frequency. In practice, some errors will be introduced in selecting the major peaks, but extreme accuracy is not essential, as accurate selection of one or more of the many wave patterns will suffice for the final analysis of phonated characters. After the frequency-stabilized wave is produced, it may be passed through certain filters, so that only predetermined frequency components of importance are present in the output, from which (either waveshape or frequency components) sets of parameters may be derived to collectively define the character of each phonated sound. It may be noted that conversion of low-pitched voice to high-pitched voice affords the use of smaller and lesser number of circuit components, for example, in filter sections, for a practical use of the apparatus.

Apparatus by which frequency transposition of the speech waves may be achieved is described in the following specification in connection with the accompanying drawings, wherein: Fig. 1 is a graphical illustration of normal speech wave, showing the manner in which wave patterns are selected; Fig. 2 is partly schematic and partly block diagram of the frequency transposer in accordance

with the invention; Fig. 3 is a schematic diagram of the major-peak detector in accordance with the invention; Fig. 4 is partly a block diagram of apparatus by which spoken words are typed phonetically; and Fig. 5 is a block diagram of a modified arrangement of fundamental frequency selector.

Speech waves

As described previously, a phonetic sound is characterized by a set of definite frequency components (disregarding the superficial frequency components introduced by the larynx) based upon a fundamental frequency. When graphical observation of these waves is made, the fundamental frequency is easily distinguished from the others by measuring the time distances between constituent waves, that have higher peak-amplitudes than the waves characterizing the original phonation. This is clearly shown in Fig. 1, wherein the time length between two major peaks represents one cycle period of the fundamental frequency. Thus, by subdividing the original speech wave at these major peaks, each wave train may be recorded separately (step by step), and reproduced several times before reproducing the next recorded wave train. The manner in which these major peaks may be selected will be described later by way of the circuit diagram given in Fig. 3, which also provides stepwise amplitude control of the original speech waves.

Frequency transposer

With reference to the diagram of Fig. 2, output signals of the major-peak detector are applied upon the waveshaper, which produces output pulses of equal widths, but at variable time intervals; depending upon the time length of each wave train. Output of this waveshaper is applied upon the saw-tooth wave generator **1**, which in turn produces saw-tooth waves of variable time lengths and of variable amplitudes; but their retrace time intervals are constant, as indicated in the drawing. Output of this saw-tooth generator is applied upon the horizontal deflecting plates of a storage tube, for example the Graphechon, as indicated in the drawing. A detailed descr'ption of this tube is given in RCA Review, vol. XII, p. 220, June 1951, by L. E. Flory et al., titled "A storage oscilloscope." The drawing in Fig. 2 however, is simplified, and electrostatic deflecting plates, rather than magnetic coils are shown; as either type may be utilized for satisfactory operation. As shown, the tube comprises on the writing side: cathode *c*–**1**; beam **1**; beam-intensity control grid G–**1**; horizontal deflecting plates *h*–**1**; and vertical deflecting plates *v*–**1**. On the reading side similar elements are provided, as: cathode *e*–**2**, beam **2**; beam-intensity control grid G–**2**; horizontal deflecting plates *h*–**2**; and vertical deflecting plates *v*–**2**. The beams **1** and **2** are projected from opposite sides upon the common target T. In order to include a brief description of the storage characteristics of the Graphechon, the following paragraph is taken from the above reference:

". . . The target or storage element consists of a fine mesh screen supporting a very thin layer of aluminum on which in turn is deposited a thin layer of insulating material. An electron beam which is called the reading beam, scans over the insulating surface and establishes a condition of potential equilibrium. Once this equilibrium is established, further scanning with no signal applied produces no signal output. If now a second beam from the opposite side of the target and of sufficient energy to penetrate the aluminum and insulating film is caused to strike the target and scan some pattern over the surface, it will upset this equilibrium condition by depositing or removing charge from the insulator. This beam is called the writing beam. This condition must then be corrected by the reading beam and the process of restoring equilibrium generates the reading signal. If the amount of charge removed or deposited by the writing beam in one scan is large compared to the amount which can be restored by the reading beam in one scan, then many scans of the reading beam will be necessary to restore the original condition, and thus many reproductions of the written signal will be obtained. This process provides the mechanism of storage utilized in the instrument to be described."

According to the reference description, and in order to adapt the storage characteristics of the Graphechon to the invention described herein, normal intensity of the reading beam is adjusted so as to obtain many scans of reading; with blanking of the beam during fly-back periods. But during the last fly-back period of reading, the beam **2** is unblanked and its intensity is increased, so as to erase the written storage for repeated writing and reading of the speech waves.

Continuing with the operation of Fig. 2, the speech wave from terminal S is applied upon the control grid G–**1**, and the output of saw-tooth generator **1** is applied upon the horizontal deflecting plates *h*–**1**, so that the beam **1** is deflected upon target T horizontally, to write and store the original input speech wave. For successive writing (recording), stepwise voltages are applied upon the vertical deflecting plates *v*–**1**, arriving from distributor **1**, which is operated stepwise sequentially by the output pulses of the waveshaper. Output voltages from different terminals of distributor **1** are mixed in the mixer **1**, so that various voltages are combined at the output to be applied upon the deflecting plates *v*–**1**. Output pulses of the waveshaper are also applied upon cathode **1** in positive polarity, so as to blank out the beam-current during retrace periods, in a manner such as practiced in television. As described previously, the horizontal writings will be of various lengths; depending upon the time interval of each succeeding wave train of the speech, as marked by the major-peaks. Thus for reading purpose, the scan period must be constant, but the amplitude of each scanning wave must be varied, so as to effect full length reading for the required frequency transposition. This is achieved by first measuring the time lengths between successive major-peaks and deriving representative static potentials, which are then utilized to modulate the amplitudes of the reading saw-tooth waves.

As distributor **1** is operated sequentially by output pulses of the waveshaper, the distributor applies steady-state positive potentials upon normally non-conductive gates **1**, **2** and **3**. These positive potentials are of such amplitudes that, the gates are sequentially driven just short of conductance. Output voltage of the saw-tooth wave generator **1** (variable in frequency and amplitude, as indicated in the block drawing) is applied upon last said gates simultaneously, so that each gate when simultaneously excited by a positive potential from distributor **1**, it admits the saw-tooth voltage to be applied upon one of the diodes V1, V2 or V3 associated therewith, and the peak voltage of the saw-tooth wave is stored in one of the condensers C1, C2 or C3, in their respective order. The steady state potentials across condensers C1, C2 and C3 are applied in positive polarity upon normally inoperative gates **4**, **5** and **6** respectively. Simultaneously, output signals of the distributor **2** are applied upon last-said gates, so that, as described in the previous manner, these gates operate one at a time sequentially, and admit the voltage-signal across an associate condenser only at the time when simultaneous positive potential arrives from the distributor **2**. Output signals of gates **4**, **5** and **6** are combined at a common output terminal as shown, and applied upon the amplitude modulator. Output of the saw-tooth generator **2** (constant in frequency, as indicated in the block drawing) is also applied upon this modulator, whereby the amplitude of each saw-tooth wave is modulated proportionally at the output of the modulator, corresponding to the original scanning time-lengths of the stored writings. Thus, the saw-tooth voltage at the output of the modulator (constant in frequency, but variable in amplitude, as indicated above block drawing of the modulator) is applied upon horizontal deflect-

ing plates *h*–2, to deflect the reading beam 2 upon the target T, whereby the originally stored wave trains are reproduced in predetermined transposed frequency regions.

In order to avoid phase distortion of the reproduced waves, the retrace periods of both saw-tooth generators 1 and 2 are adjusted to be the same. Such control is simple, and has been practiced in the art of electronics. In the drawing however, the saw-tooth generator 2 is shown controlled by output pulses of the constant frequency pulse generator; the pulse-width of which is adjusted to be equal to the output pulse of the waveshaper. Output of the constant frequency pulse generator is applied upon cathode *c*–2 in positive polarity, so that the reading beam 2 is blanked out during retracing periods. However, at the time when the reading beam 2 is to be shifted in vertical position to read the next written storage, then at the time of retracing the reading beam is unblanked, and the potential upon grid G–2 increased to an extent, as to increase the beam current for erasing the written storage on that horizontal position of target T. Unblanking and increasing intensity of the beam 2, is achieved by simultaneously applying output pulse of the pulse generator upon the control grid G–2 in positive polarity; through normally inoperative gate 7 and the phase inverter; at the time when the pulse-prolonger is operating.

In order to synchronize operational sequence between distributors 1 and 2, the former being operated ahead of the latter, operation of the distributor 2 is controlled by the trigger circuits of distributor 1. For example, first of the three triggers of distributor 1 operates the third of three triggers of distributor 2; second of the three triggers of the former operates the first of the three triggers of the latter; etc. Output signals of distributor 2 are then combined in mixer 2, and applied upon the vertical deflecting plates *v*–2, to shift the reading beam 2 to the next written storage position.

In order to avoid cross fire between writing and reading beams; due to differences in time periods of the two scannings, a delay control is provided for the reading beam. To achieve this time delay, output operating-signals of the distributor 1 are individually applied upon distributor 2 through the normally inoperative gates 8, 9 and 10. During transition period of any of the triggers operating in distributor 1, a voltage pulse may be derived. Hence, each time the voltage in mixer 1 changes, a voltage pulse is derived, for example, by the small condenser *c*, and applied upon the block of pulse-prolonger, which prolongs the input pulse not greater than the time interval between pulses of the pulse generator. The gate 7 is normally so adjusted that, it operates only when simultaneous positive potentials arrive from the pulse generator, and pulse-prolonger. Thus, at the time when a trigger in distributor 1 operates, a pulse is applied upon the pulse-prolonger, which in turn applies a positive potential upon gate 7. This gate remains idle until it receives a simultaneous positive potential from the pulse generator, at which time, a short pulse passes therethrough, and applied in positive polarity (passing through the phase inverter) upon the control grid G–2, to increase the current-intensity of beam 2, for erasing the recorded signal during fly-back period. In order to prevent shifting of the beam's vertical position during this retrace, output pulse of the phase inverter (or gate 7) is delayed by the delay block, and applied in positive polarity upon the gates 8, 9 and 10 simultaneously, whereby the distributor 2 operates and shifts the vertical position of the beam 2, by a static voltage passing through mixer 2, for scanning (reading) the next train of written speech wave. Simultaneously, the distributor 2 applies a positive voltage upon the control grid of a discharger tube (V4, V5 or V6) associated with the distributor-trigger that had been operated previously, and causes that tube to conduct from previous non-conductive state to discharge the stored potential of its associated condenser (C1, C2 or C3), so as to prepare it for a repeated charge; when the distributor 1 returns its cyclic operation to that stage. When operating time of the triggers in distributor 2 is equal to one retrace period, then the delay block is not necessary.

In reference to removal of the stored writing from target T, a high frequency noise will be introduced at the output terminal S' of the reproduced speech, at that time. While such noise will not affect the final analysis of speech waves (also because such high frequency may be easily filtered out from the output), a negative pulse from gate 7 as shown (or from another section) may be applied at terminal S', if found necessary. The circuitry in block diagrams, such as: distributors (which may be other than triggers, as indicated in the drawing, such as a cathode-beam sweeping over mutually insulated targets); gates; phase inverter; pulse generator; waveshaper (which may comprise a trigger circuit operated by the major-peak signal; the trigger in turn producing output pulses of the predetermined width); saw-tooth wave generators; modulator; pulse prolonger; and mixers are well known in the art of electronics, and therefore, detailed explanation is not found necessary herein. Similarly, the Graphechon employed in the diagram is given only as an example, as there are different methods and apparatus for writing and reading electrical signals. For example, the image pick-up devices, such as the Orthicon, Vidicon etc., used in television, are storage devices; it is a matter of how much storage is required for a particular purpose. Accordingly, other types of storage devices, having the required storage characteristics may be utilized in connection with the present invention. For the purpose utilized herein, the storage device is not required to be of highly precise type, and satisfactory operation will be obtained without critical adjustments, since during each recording of a wave-pattern across horizontal line, there may be at most 20 wave-cycles; and only few lines are required for cyclic recording. In Fig. 2, three storage condensers; chargers; and dischargers are shown. It is therefore assumed that the distributors are also arranged for tri-signal cyclic operation. However, these are exemplary, and the number of such stages may be otherwise.

Printing apparatus

As described in the foregoing, and in reliance with previous tests, it had been stated that each phonetic sound is characterized by a definite set of frequency components; each having different amplitude with respect to the other. In the original speech, these frequency components vary in range as the fundamental (pitch) frequencies vary. But since the transposed and reproduced speech wave at the output terminal S' will contain only a single fundamental frequency, there will be substantially only a single set of frequency components of importance, which in different combinations characterize different phonetic sounds of the original speech. Thus, at the output terminal S', a number of tuned circuits in resonance with the frequencies of importance, such as f_1, f_2, f_3, f_4 and f_5, are connected as shown in Fig. 4, from rectified outputs of which are derived discrete signals to operate a printing device, such as a modified electric typewriter. These discrete signals are produced across output resistance R0, by coupling different combinations of resistances R1 to R5 respectively, to a common voltage supply B0. The different signals across R0 are distinguished by differences in the quantum size, assigned to each of the cumulative combinations. That is, the various quanta size, or steps are so related that the total corresponding to any particular combination of relay operation is distinctly different from the total corresponding to any other combination that might occur. The circuit arrangement as shown may be referred to patent application Serial No. 114,446, filed September 7, 1949, by R. E. McCoy and myself now U. S. Patent No. 2,618,706,

November 18, 1952. For clarity however, the following description is included herein.

When relay S*n* (in the *n*th branch) is operated, and its armature-contact is depressed (down), it connects a corresponding resistor R*n* between battery B0 and load resistance R0. A current flows through the closed circuit, producing across R0 a voltage proportional to that current and the resistance R*n*. If several of the relays are operated and their armature-contacts are depressed simultaneously, the resulting output voltage (as a fraction of the voltage of battery B0) is:

$$(\sum_m 1/R_m)/(1/R0-\sum_n 1/R_n)$$

where the *m*-summation in the numerator is taken for only the resistances controlled by those armature-contacts which happen to be down, while the *n*-summation in the denominator is taken for the resistances in all branches, regardless of contact position.

By suitable choice of resistances, the change of output voltage controlled by each armature-contact can be made any desired fraction of the voltage from common source B0. For example, in a system with a total of N branches, if the resistance switched by the *n*th relay S*n* is

$$R_n=2^{1-n}R0 \text{ (for } n=1, 2, 3 \ldots, N)$$

then operating S*n* will add to the output voltage a fraction 2^{n-1-N} of the voltage of common source B0. For the Nth branch, this fraction is ½; and the maximum output, with all relay-armature contacts down, is a fraction $(1-2^{-N})$ slightly less than unity. If R0 were then removed or changed to an open circuit, keeping the other resistances unchanged, the maximum output could be increased to the full value of the voltage from common source B0, and all other output voltages would be increased in the same proportion. Thus, it may be seen that any combination of relay operation will produce a different step of voltage across R0, that is distinctly different from the total corresponding to other combinations, and each of these voltage steps represents a phonetic character in the original speech waves.

To detect these signals, the voltage across R0 is applied upon the electrostatic deflecting plates *d* of cathode ray tube CRT, wherein, the beam *e* is deflected across mutually insulated targets, as indicated in the drawing. Output signals of these targets are then applied independently (amplified if necessary) upon pre-arranged solenoids of a letter-typing device, for example, solenoids of a modified electric typewriter, to operate appropriate keys for typing the phonetics contained in the original speech.

In order to avoid cross fire at time of the beam *e* shifting from one target to another, for exciting different keys of the phonetic printer, the changing voltage across R0 is differentiated by a small condenser *c*, and applied upon the beam-intensity control grid G, in negative polarity, so as to extinguish the beam while changing its horizontal position. In actual operation however, blanking of the beam may not be necessary, since the time of beam movement from one target to another is very short for mechanical key operation. However, the duration of a phonetic sound is long enough to effect operation of the typing key. For example, a moderate rate of talking will give about 400 phonetic letters per minute, and an ordinary teletypewriter is capable of operating at a high speed of about 600 characters per minute. Present commercial devices are capable of performing 3000 characters per minute.

To effect word separation between typed letters, the original speech wave at terminal O is full-wave rectified through diodes V7 and V8 across R-C circuit. The time constant of this circuit is so adjusted that, the charge across the condenser is dissipated in about the shortest time interval that the speaker may pose between spoken words. Thus during quiescent periods the voltage across R-C circuit is minimum, and will not pass through the normally inoperative gate 11. But when the speech wave appears, a voltage of substantial amplitude will appear across R-C circuit, and pass through gate 11, the output of which will operate the space key of the printer.

In reference to previous tests, the number of frequency components characterizing a phonetic sound had been determined, and accordingly, few of the voiced phonetic sounds have been successfully produced by means of sets of synthetic wave production. Such tests however have been subject to wide frequency variations, and almost impossible to provide thorough information, due to the widely varying fundamental (pitch) frequencies. As described in the foregoing, these variables are eliminated in the present invention, and therefore, the number of resonant circuits f_1, f_2 etc., in Fig. 4 is substantially fixed. With the transposed waves arriving at terminal S', the number and frequency components of importance will have to be re-established, which may be done very easily due to the limited frequency range. As a good approximation however, the highest number of frequency components of importance contained in any phonetic sound will not be higher than the five different frequencies shown in the drawing, which by different combinations will effect a total of 32 voltage-steps across R0, and accordingly, 32 targets in the cathode ray tube are indicated for the operation and printing of 32 different letter-characters.

When the amplitude differences of these frequency components are also to be included in the process of translation into printed phonetics, then the armature contacts of relays S*n* may have more than two contacts, each in series with a resistance as described previously, so that the distance of armature movement, depending upon the output power of any one of the rectifiers, will effect the correct voltage step across R0. Alternatively, more than one relay (each responsive to different applied voltage) may be connected in parallel (or otherwise arranged) at the output of each rectifier, so that when output voltage of the rectifier is low, only one relay will operate and cause a certain voltage change across R0, while when output voltage is high, more than one relay will operate; effecting different voltage step across R0.

Major-peak detector and stepwise amplitude control

In order to eliminate widely varying amplitudes of the original speech waves, a stepwise amplitude control is included in conjunction with the major-peak detector. In Fig. 3, the original speech wave arriving at terminal O, is applied upon the control grid of modulator tube V9, which amplifies the input signal in the plate tank circuit comprising coil L1. The signal-gain in output circuit L1 is controlled by varying the screen voltage of V9, by the driver tube V10, which, when changed in anode conductance by a change in grid potential, causes corresponding voltage change in resistance R6, applied directly upon the screen of modulator tube V9. The magnitude of speech signal at input terminal O is normally adjusted higher than the desired value, so that when the grid of tube V9 is driven highly positive at a major peak, an increasingly positive voltage from the inverted terminal of tank coil L1 is transmitted to the coil L2, until this voltage is equal to the voltage of B1; beyond which, the anode potentials of diodes V11 and V12 become positive with respect to their cathode potentials and start conducting; with resultant increase in charges across condensers C4 and C5. The positive voltage across condenser C4 is applied upon the control grid of driver tube V10, causing magnified voltage drop across resistance R6, and corresponding gain-drop in modulator tube V9. The extent in which this gain-control is achieved is dependent upon the voltage gain of driver tube V10. Although high gain control is not essential, the circuit may be modified for high gain control, such as employed in radio receiving circuits, and other practices.

The major peak pulse is derived from the charge across condenser C5. That is, once the condenser C4 is charged to the peak, its charge remains constant until another major peak of the speech signal arrives at the input ter-

minal O. Whereas, in the process of charging condenser C5 there is produced an output signal which is amplified by the pulse amplifier, the output of which represents the major peak. In order that the condenser C5 may be recharged during the following incoming major peak, output positive pulse of the pulse amplifier is applied upon the control grid of discharger tube V13, through the delay circuit; causing V13 to conduct for a short time and discharge condenser C5 by a predetermined quantity; which is determined by the series resistance R7. Thus, when the following major peak arrives from the input terminal O, the condenser C5 is recharged to the previous peak value, and causes another pulse at the output of the pulse amplifier. To ensure that the gain of speech signal has not fallen below the controlled value, the output positive pulse of the pulse amplifier is applied upon the control grid of normally non-conductive discharger tube V14, which starts conducting when the newly applied voltage from across coil L2 is less than the previously stored voltage in condenser C4; causing a corresponding increase in G_m of modulator tube V9, until the maximum control of the output gain is re-established.

The purpose of major peak detector given in Fig. 3 is to subdivide the speech wave in successive wave-trains at one cycle periods of the fundamental frequencies, so that they may be reproduced in different time periods for frequency transposition. Accordingly, the original speech may be subdivided at other points than the major-peaks, as long as the time periods of these divisions are substantially the same as before.

Modified arrangement of speech subdivider

Fig. 5 is a block diagram of another form of subdividing the speech waves. In this arrangement, resonant circuits are provided to select the fundamental frequencies. In the average the fundamental frequencies in speech waves range from 90 to 300 cycles per second. During an articulated phonetic sound, the lowest frequency contained therein represents the fundamental frequency, and all other frequency components must be at least twice as high or higher than the fundamental frequency. To select only the fundamental frequency during a phonated sound, there are arranged five resonant circuits, tuned in steps to frequencies as indicated in the graph of Fig. 5. Each of these resonant circuits is represented by the blocks f_a to f_e, respectively. Outputs of these resonant circuits are full-wave rectified in blocks *a* to *e*, respectively, and outputs of the rectifiers are applied to the inputs of gates *a* to *e*, respectively, to control admittance of the selected fundamental frequencies. Frequency f_a represents the lowest resonant frequency, and f_e represents the highest fundamental frequency. Since at any given time the fundamental frequency represents the lowest frequency in speech waves, operation of gates *a* to *e* is such that, a rectified voltage appearing at the output of a rectifier connected to the block of the lowest frequency, cuts off the operation of all gates connected to next higher resonant circuits, for example, output negative voltage of rectifier *a* cuts off the operation of gates *b* to *e*; the negative output voltage of rectifier *b* cuts off the operation of gates *c* to *e;* etc. Each succeeding rectifier also applies a negative potential upon a preceding gate, for example, *b* to *a*; *c* to *b*, etc.; but in lower amplitude than one rectifier to a succeeding gate, so that an immediately succeeding rectifier is also capable of cutting off the operation of a preceding gate; but at a later time period than the preceding rectifier is capable of cutting off the operation of the succeeding gate. Such cross-application of negative potentials in different amplitudes provides better switching from one step to the other, due to the overlapping response curves of the resonant circuits. Thus, when the fundamental frequency in speech wave at a given time is at the crossing point (as shown in the graph), the gate passing the lower frequency becomes operative; while when the fundamental crosses closer to the higher frequency, the gate passing the higher frequency becomes operative.

Outputs of the gates *a* to *e* are combined in the mixer, which in turn are applied to the major-peak detector (similar to the arrangement given in Fig. 3), whereby subdividing (of speech waves) pulses at intervals of one cycle periods of the selected fundamental frequency may be obtained at its output. Since output of the mixer contains only the fundamental frequency at a given time, subdividing pulses may also be obtained by circuits other than the major-peak detector shown. However, the major-peak detector provides automatic gain control of the speech wave. In Fig. 5, the resonant circuits as shown, may be arranged differently; having different response curves, as best suited for the purpose. Similarly, the gates *a* to *e* may be inserted before the resonant circuits, f_a to f_e, inclusive; or in combination.

It will be obvious from the general principles herein disclosed that, numerous substitutions of parts, adaptations and modifications are possible without departing from the spirit and scope thereof.

What I claim is:

1. Apparatus for transposing unknown frequency components of importance in speech waves to regions of substantially known frequencies which comprises: means to produce speech waves, a frequency-detector and means therefor to detect the fundamental frequencies contained in said speech waves, means to produce output pulses at one cycle periods of detected waves of said frequency-detector, a wave-recorder and means therefor to record said speech waves, a first distributor, means to apply said pulses upon the distributor so as to operate it stepwise cyclically, means to apply output signals of said distributor upon the wave-recorder so as to shift the recording positions step by step, whereby to obtain independently sequential recordings of the speech wave in wave-trains, means to measure the time lengths of the wave-trains at time of recordings and means therefor to derive therefrom representative static signals, a series of storage means, means to store the static signals in the storage means sequentially under control of said distributor signals, means to scan the recorded wave-trains in predetermined constant time intervals so as to reproduce previously recorded waves while succeeding wave-trains are being recorded, means to apply said stored signals upon last said scanning means in orderly sequence so as to control the scanning amplitudes in agreement with the original scanned distances of the recorded wave-trains, a second distributor, means to apply output signals of the first distributor upon the second distributor for stepwise operation in a manner that when one recording position is changed to another the scanning position is changed forward after completing its previous scanning, whereby the scanning follows the recordings stepwise in proper sequence, means to apply output signals of the second distributor upon said storage condensers in lagging sequence to discharge same, whereby they may be re-charged at time of repeated recordings, and means to erase previously scanned recordings, whereby repeated recordings are provided for; said scanned (reproduced) wave-trains in constant time periods representing said frequency-transposed waves.

2. As set forth in claim 1, wherein said fundamental frequency-detector comprises a first electron tube having anode; cathode; and multi electron-control elements, a first impedance associated with said tube, means to apply the original speech wave upon one of said control elements, whereby the applied waves appear proportionally in the first impedance, first rectifier; first storage condenser; and means therefor to store peak voltage of the waves appearing in the last said impedance in the first condenser through the first rectifier, means to apply the output of said stored voltage upon the other control element of the first electronic tube in pre-phased polarity and approximately predetermined amplitude, whereby to adjust the gain of said tube to substantially a predetermined

level, and thereby effecting substantially a predetermined peak potential in the first impedance, second storage condenser; second impedance and first bias supply connected in series; means to apply the voltage appearing in the first impedance to the second impedance; the magnitude of said bias being so adjusted that the second condenser is charged proportionally only at near the peak of the arrival of said major-peak of said speech wave upon said control element of the first tube, whereby the output signal of last said charge represents said major-peak, a second electron tube having anode; cathode; and electron-control element, means to connect last said tube in series with the first condenser; second impedance; and the first bias, a second bias supply upon the control element of the second electron tube, whereby said tube is normally rendered non-conductive, means to apply the output signal of the second condenser upon control element of the second electron tube for a short time interval, so as to render it conductive and discharge the first condenser by an appreciable amount, whereby last said condenser may be re-charged anew for stepwise repeated gain-control of the first electron tube, a third electron tube having anode; cathode; and electron-control element, means to connect the anode-cathode of the third tube across the second condenser, a third bias supply upon control element of the third tube to render it normally non-conductive, and means to apply the output signal of the second condenser upon control element of the third tube to render it conductive and discharge the second condenser by a predetermined amount, whereby upon the arrival of the succeeding major-peak last said condenser is re-charged to produce an output signal representative of said major-peak.

3. As set forth in claim 1, which includes apparatus for detecting and translating said transposed waves into printed letters representing the phonetics contained in the original speech waves, which comprises: a set of independent frequency-selective circuits responsive to frequency components of importance in said transposed speech waves, independent relay or relays having armature contact or contacts at the output of each of said circuits said relays being operative by the outputs of their associate frequency-responsive circuits, a voltage source, coupling means between the voltage source and a common output circuit; said coupling means having a plurality of impedance-branches controlled by the armature contacts of said relays; each impedance designed to produce a different increment of voltage at the said common output terminals when the corresponding contact is operated, in a sense that any combinations of said contacts operated at a time will produce distinctly different increments representing significantly different phonetics contained in the original speech waves, means to detect last said incremental voltages, whereby to obtain independent output signals, a printing or typing apparatus having independent input terminals for each letter character, means to apply aforesaid signals to last said terminals, whereby to translate speech waves to printed letters, and in combination, means to measure the average value of the original speech waves, and means to shift the spacing of said printed or typed letters, whereby substantially word separation may be provided.

4. As set forth in claim 1, which includes apparatus for detecting and translating said transposed waves into printed letters representing the phonetics contained in the original speech waves, which comprises: a set of independent frequency-selective circuits responsive to frequency components of importance in said transposed speech waves, independent relay or relays having armature contact or contacts at the output of each of said circuits; said relays being operative by the outputs of their associate frequency-responsive circuits, a voltage source, coupling means between the voltage source and a common output circuit; said coupling means having a plurality of impedance-branches controlled by the armature contacts of said relays; each impedance designed to produce a different increment of voltage at the said common output terminals when the corresponding contact is operated, in a sense that any combination of said contacts operated at a time will produce distinctly different increments representing significantly different phonetics contained in the original speech waves, means to detect last said incremental voltages, whereby to obtain indepedent output signals; said detector comprising a cathode discharge device; having electron beam; beam deflecting means; and a plurality of mutually insulated targets in the path of said beam, means to apply last said incremental voltages upon last said deflecting means, whereby to deflect the beam upon said targets; said targets so positioned as each to intercept the beam at the angular deflected position corresponding to any of said incremental voltages, a phonetic printing or typing apparatus having independent solenoids for each typing key, means to derive output signals from each of said targets independently and means therefor to apply same in predetermined arrangement upon said solenoids so as to operate same for printing of phonetic characters.

5. In speech waves where all frequency components of importance representing phonetic sounds are based on unknown fundamental frequencies, the system of transposing all unknown frequency components to known frequency regions based on a reference fundamental frequency, which comprises means for propagating speech waves, means for assigning a reference fundamental frequency, means for selecting the variable fundamental frequency components of the propagated waves, frequency-measuring means; and means therefor; for measuring the differences between said selected fundamental frequencies and said reference frequency, and according to last said differences means for shifting substantially all frequency components of the propagated waves to frequency regions where all varying fundamental frequencies are substantially equal to said reference frequency, thereby transposing all frequency components of the speech waves to frequency regions where they are all based substantially on a single reference (fundamental) frequency.

6. The system as set forth in claim 5, which includes means for selecting from said frequency transposed speech waves sets of frequency components any combination of which collectively represents a distinct phonetic sound, means for deriving distinct signals from said sets of selected components, and means for translating said sets of distinct signals into visible intelligible indicia representative of said phonetic sounds.

7. The system as set forth in claim 5, which includes means for translating said transposed waves into visible intelligible indicia representative of the original phonated sounds, which comprises means for selecting sets of frequency components of importance from said transposed waves, means for deriving significantly different quantities from those sets of frequency components, means for totalizing said quantities in a manner as to obtain a different step of quanta for each set of said totalized quantities; each of said quanta representing significantly a different phonetic sound in the original speech waves, means for detecting said steps of quanta, apparatus for printing intelligible indicia, and means for applying said detected steps of quanta to operate last said apparatus for printing intelligible indicia representative of the original phonations in the speech waves.

8. In speech waves where all frequency components of importance are based on unknown fundamental frequencies, the system of transposing all unknown frequencies substantially to known frequency regions based on a standard fundamental frequency, which comprises means for producing speech waves, means for assigning a standard fundamental frequency, fundamental-frequency detector for detecting the fundamentals of the produced speech waves; and means therefor for dividing the produced speech waves into substantially one cycle

portions of the detected fundamentals, means for recording last-named divided portions at some normal speed, means for measuring the time differences of recordings of said divided portions with that of one cycle periods of the standard frequency, and reproducing-means under control of said measurements for reproducing said recorded portions equal in time of one cycle periods of the standard frequency, thereby shifting all frequency components of the speech to regions based on a standard fundamental frequency.

9. The system as set forth in claim 8, which includes means for selecting from said frequency transposed speech waves substantially those frequency components that collectively compose the original phonetic sounds, and means for translating last-said collective components into visible intelligible indicia representative of the original phonetic sounds.

10. In speech waves where the phonetic characters are composed of certain combinations of frequency components relative to a fundamental, and wherein the frequency position of said fundamental is neither predictable nor always selectable by resonance means, the system of transposing the instantaneous frequency components of the speech waves to regions where said combinations will substantially always have standard frequency relations to a reference fundamental frequency which comprises means for producing speech waves, a major-peak detector and means therefor for detecting the major peaks of the produced speech waves, means for assigning a reference fundamental frequency, time-measuring means and means therefor for measuring the time periods between the detected major peaks as representative one cycle portions of the unknown fundamental; and for deriving representative quantities corresponding proportionally to the differences of time between the measured and one cycle time periods of said reference frequency, and variably controlled frequency-shifting means for shifting the instantaneous frequency components of the produced speech waves at varying degrees under stepwise control substantially proportional to said quantities, whereby substantially equalizing the time periods between said major peaks with that of one cycle portions of said reference frequency, and thereby substantially standardizing the frequency relations of said combinations relative to a reference fundamental frequency.

11. The system as set forth in claim 10, which includes means for selecting from said frequency transposed speech waves those combinations of frequency components that compose phonetic characters, and means for translating these selected combinations into discrete signals representative of said phonetic characters.

12. The system as set forth in claim 10, which includes means for selecting from said frequency transposed speech waves those combinations of frequency components that compose phonetic characters, means for translating these selected combinations into discrete signals, and means for translating these discrete signals into visible intelligible indicia representative of said phonetic characters.

13. The system as set forth in claim 10, wherein said major-peak detector comprises a receptive means and an amplifier means of said speech waves, a bias source and adjustment means in the amplifier means for causing the oppositely polarized speech waves as if they were varying unidirectionally from the bias source, and a peak detector means for detecting substantially the highest peaks of the speech waves as measured from said bias source.

14. The system as set forth in claim 10, wherein said major-peak detector comprises a receptive means and an amplifier means of said speech waves, a bias source and adjustment means in the amplifier means for causing the oppositely polarized speech waves as if they were varying unidirectionally from the bias source, a peak detector means for detecting substantially the highest peaks of the speech waves as measured from said bias source, and a polarizing means for switching the polarity of the speech waves in said amplifier, whereby selecting the polarity of waves that contain most distinguishable peaks for said major-peak detection.

15. In speech waves where the phonetic characters are composed of certain combinations of frequency components relative to a fundamental, and wherein the frequency positions of said fundamental is neither predictable nor always responsive to resonance means, the system of transposing the instantaneous frequency components of the speech waves to regions where said combinations will substantially always have standard frequency relations to a reference fundamental frequency which comprises means for producing speech waves, a major-peak detector and means therefor for detecting the major peaks of the produced speech waves, means for assigning a reference fundamental frequency, time-measuring means and means therefor for measuring the time periods between the detected major peaks as representative one cycle portions of the unknown fundamental; and for deriving representative quantities proportionally corresponding to the differences of time between the measured and one cycle time-periods of said reference frequency, recording means and means therefor for recording the speech waves occurring between said major peaks independently step by step at some normal speed, reproducing means and means therefor for reproducing the recorded speech waves, and control means for varying the reproduction speed by said quantities step by step, whereby equalizing the time lengths between the reproduced major peaks to that of one cycle portions of the reference frequency by stretching and compressing, and thereby effecting said frequency transposition.

References Cited in the file of this patent

UNITED STATES PATENTS

2,195,081	Dudley	Mar. 26, 1940
2,375,044	Skellett	May 1, 1945
2,540,660	Dreyfus	Feb. 6, 1951
2,570,858	Rajchman	Oct. 9, 1951
2,575,910	Mathes	Nov. 20, 1951
2,593,694	Peterson	Apr. 22, 1952